Connective Tissue Research, 44(Suppl. 1): 33–40, 2003

0300-8207/02 $12.00 + .00
DOI: 10.1080/03008200390152061

Taylor & Francis
healthsciences

Six Genes Expressed in Bones and Teeth Encode the Current Members of the SIBLING Family of Proteins

Larry W. Fisher[1] and Neal S. Fedarko[1,2]
[1]*Matrix Biochemistry Unit, Craniofacial & Skeletal Diseases Branch, NIDCR, NIH, HHS, Bethesda, Maryland, USA*
[2]*Division of Geriatrics, Department of Medicine, Johns Hopkins University, Baltimore, Maryland, USA*

Bone sialoprotein (BSP), dentin matrix protein 1 (DMP1), dentin sialophosphoprotein (DSPP), enamelin (ENAM), matrix extracellular phosphoglycoprotein (MEPE), and osteopontin (OPN) are glycophosphoproteins expressed in bones and/or teeth. Direct comparison of their amino acid sequences do not suggest that they belong to a single genetic family, but a detailed analysis of their chromosomal location and gene structure does. Analysis of human brain mRNA by RT-PCR has led to the discovery of two additional exons thereby making it more convincing that MEPE is a member of the SIBLING (Small Integrin-Binding LIgand, N-linked Glycoprotein) family. We propose that the members of this SIBLING family are extended, flexible proteins in solution that can facilitate the formation of a number of different complexes. For example, OPN can bridge complement Factor H to either an RGD-dependent integrin or to CD44 forming a membrane-bound complex that actively suppresses the alternate complement pathway. Two possible mechanisms for inhibiting the lytic pathway of alternate complement are presented.

Keywords Bone Sialoprotein, Complement, Integrin-Binding, Osteopontin, SIBLING.

INTRODUCTION

Studies of the extracellular matrices of bones and teeth have a long and rich history. Ages before written scientific literature it was noticed that bones treated with a weak acid such as vinegar changed into supple structures that looked and behaved much like the skin, tendons, ligaments, and other soft tissue elements. The application of first modern protein biochemistry methods and then more recently, molecular biological approaches, have led us to understand much of the biology and mechanical properties of the underlying collagen scaffolding that constitutes the vast majority of the matrix of bones and dentin. Indeed, the vast majority of the genetic diseases of bones and teeth whose mutations are known are mutations in the various members of the collagen family [1].

Still, the age-old question of why bones and teeth mineralize whereas the nearly identical skin, tendon, and ligaments do not remained unanswered and the quest to understand this intriguing process was undertaken by a number of laboratories around the world. Many of the most successful of these laboratories started about 30 years ago, during the ascendancy of the idea that one gene leads to one protein and that most biological functions can be satisfactorily explained by finding and describing the protein that performs such a function. In the 1960s and 1970s a number of laboratories around the world began the search for the holy grail of "the proteins" that nucleate and/or control the growth of hydroxyapatite crystals in calcified cartilage, bone, dentin, and enamel. Logically the candidate gene products would be acidic proteins, possibly phosphoproteins, with a strong affinity for hydroxyapatite. Most investigators thought that these proteins would likely (but not necessarily) be entrapped within the mineralizing matrix and therefore be released from the mature tissues by demineralization. Only a handful of proteins were found in relative abundance in the mineralized compartment of bones and teeth and these are now the proteins whose names (often based on the names used for the cDNA sequences) we read many times in the literature: osteocalcin (OCN), osteonectin (ON), osteopontin (OPN, also known as SPP1 and Eta1), bone sialoprotein (BSP), matrix gla protein (MGP), decorin (DCN), biglycan (BGN), dentin matrix protein 1 (DMP1), dentin sialophosphoprotein (DSPP), enamelin (ENAM), amelogenin (AMEL), and others.

At least in the case of bone, the mice with null mutations in the various candidate genes do not show abject failure of mineralization. Bones are larger or smaller, thinner or thicker, perfectly

Received 9 November 2001; revised 1 February 2002; accepted 7 February 2002.

Address correspondence to Dr. Larry Fisher, Room 228, Building 30, NIDCR-NIH, Bethesda, MD, 20892-4320, USA. E-mail: LFISHER@dir.nidcr.nih.gov

formed or somewhat distorted, but all single knockout mice have grossly functional, mineralized bones. While the usual descriptions of redundancy of genes have been offered to explain why the proteins of such logical promise have disappointed us within the context of mineralization, these results also set us free to discover other functions for these curious proteins. In this paper, we argue that five of these proteins, BSP [2], DMP1 [3], DSPP [4], MEPE (matrix extracellular phosphoglycoprotein [5], also known in the rat as OF45 and osteoregulin [6], and OPN [7] are likely members of a family of proteins we have called SIBLING for Small Integrin-Binding LIgand, N-linked Glycoprotein. A sixth protein, ENAM [8], may be a more distant member of the family. Earlier work had suggested that at least four and possibly five of these genes were members of this family [9]. While we argue that at least three of the SIBLINGs do function in complement, the SIBLING name includes only biochemical descriptions and not higher biological functions because the ultimate functions of all the members of the family are not known at this time.

RESULTS AND DISCUSSION

The most common method of defining the relationship among a group of proteins is to compare their linear amino acid sequences. For the six proteins currently proposed to be in the SIBLING family, this approach does not yield satisfying results. Figure 1 illustrates this point by showing the comparison of three random pairing of members of the family using the "Compare" program of the GCG group (Accelrys Inc.). Two proteins with strong homologies will have a distinct diagonal line such as that seen for the comparison of two small proteoglycans, decorin and biglycan (Figure 1D). The comparisons of the various SIBLINGs show no such strong diagonal lines implying poor homologies at the amino acid level. Indeed, the results are probably no better than one would expect from a comparison of any two random protein sequences. There are a few short regions that are conserved among members of the family including the completely conserved integrin-binding tripeptide, RGD, and NXS/T motif for N-linked oligosaccharides as well as a number of casein kinase II-type phosphorylation sites.

Directly comparing the locations of these short sequences within the primary protein sequences in their entirety, however, does little to make a case for significant homology among the different proteins. The overall chemical properties of these proteins also seem to suggest that they are not related. For example BSP, DMP1, DSPP, and OPN are all acidic with predicted isoelectric points of 3.4 to 4.3 (without post-translational modifications) whereas ENAM is neutral and MEPE strongly basic (pI = 9.2). Even within the acid members, BSP is glutamic acid-rich and others are either aspartic acid-rich or a mix of the two. However, several important points of similarity within the genetic structures of these 6 proteins permit us to propose that they are all a result of an ancient gene duplication and subsequent divergence.

Five of the six genes are located within a contiguous region of chromosome 4q21.3 (Figure 2). The Human Genome Project has not completed this portion of chromosome 4 so the exact distances between the genes are not known, but currently five are thought to be within an estimated 750,000 base pair segment and four of those within a single 250,000 bp domain. The reader should be cautioned, however, that this region of chromosome 4 is based on incomplete sequences and the final orientations and locations of the gene will not be completely known until all the sequencing is complete. It is clear that the most similar five SIBLINGs are very closely spaced and this makes for a significant problem in producing double knockout mice. The typical method of producing double knockout mice by cross-breeding single KO mice cannot easily be done. The genotypes of hundreds to thousands of offspring of the breeding pairs would have to be checked to hope to detect a single cross-over event between genes as closely spaced as the SIBLING genes. Notice that MEPE, probably the most different member of the family, is located in the center of this close cluster of genes.

As of the writing of this report (2001), there is only one know gene between the DSPP and OPN (usually listed as SPP1 for secreted phosphoprotein 1 within the Human Genome Project) except the other SIBLINGs. The ABCG2 (for ATP-Binding Cassette Transporter, subfamily G, member 2, also known as Breast Cancer/Mitoxantrone Resistance Protein (BCRP/MXR) [10] gene was until recently mapped to a position outside the SIBLING cluster but the most recent build has it between OPN and MEPE. ABCG2 is structurally unrelated to the integrin-binding proteins but curiously is upregulated in placenta and many tumors, much like BSP and OPN [10–12]. The gene for enamelin, ENAM, also is on human chromosome 4 and is currently being assigned a position much closer to the centromere, ~4q13, but it is possible that this location will be refined at a later date.

A clustering of genes within a single chromosome alone, of course, is not justification for defining a family of gene products. The next evidence for the grouping of the proteins is the similarity of their intron-exon boundaries and the biochemical similarities of their corresponding exons. First, we report on our recent findings of additional exons for human MEPE gene. We performed PCR on reverse-transcribed mRNA cDNA from human brain (Invitrogen, Human Tissue Panel #1) using an oligonucleotide pair derived from the beginning and end of the coding region as defined by the original description of the MEPE mRNA [5]. The oligonucleotide pair incorporated restriction enzyme sites for subcloning into an adenovirus shuttle vector. After gel purification of the band of the approximate expected size, the PCR products were subcloned into the vector and 30 cDNA clones purified. Nine of the clones were identical to the original sequence previously described by Rowe et al. [5].

A standard BLAST analysis of this sequence against the human genome database identified three exons within the sequenced genome and a short section of 54 basepairs that have not yet been identified in the project but is very likely to represent a single exon, exon 3, rather than two or more exons. The other clones all had longer sequences within them that when compared with the human genomic sequences, were found to represent two additional exons, 4 and 5. Exon 4 was homologous

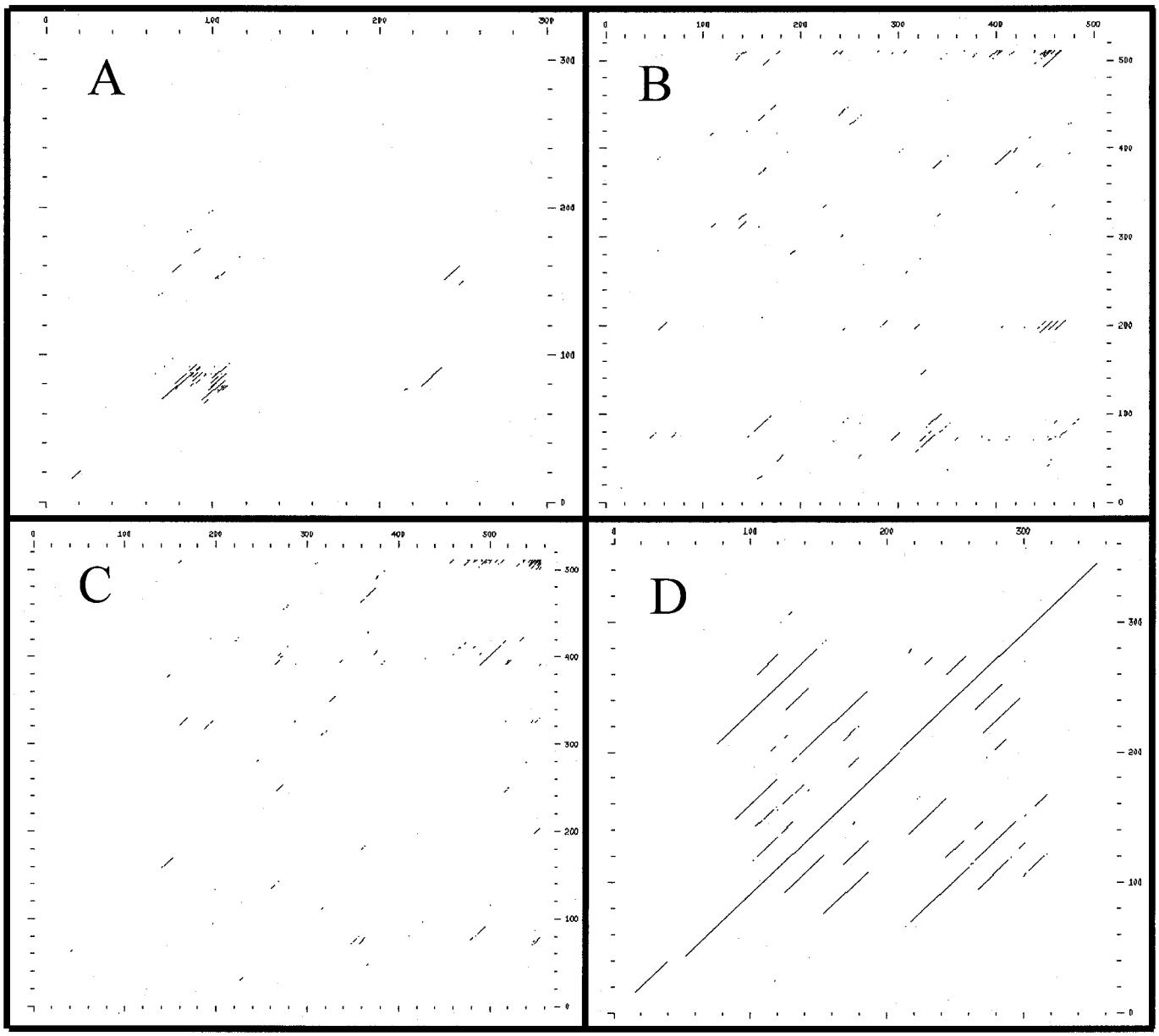

Figure 1. Graphical comparisons of paired human protein sequences using the "Compare" program of the Genetics Computer Group. (A) OPN vs. BSP; (B) MEPE vs. DSPP (with only two copies of the carboxyterminal phosphorylation repeats); (C) DMP1 vs. MEPE; and (D) Biglycan vs. Decorin. Protein pairs showing strong homology have the clear mid-panel diagonal line like that seen in the comparison of two closely related proteins, Biglycan and Decorin (panel D). Short lines off the mid-panel diagonal result from internal repeats that are shared. Notice that the SIBLING comparisons have no mid-panel diagonal lines and only short homologies throughout their lengths. This illustrates that the SIBLINGs are not very homologous at the primary sequence level.

to the extra sequences observed for monkey brain MEPE in GenBank accession number AB046056 (Osada, N et al., GenBank accession number AB046056, otherwise unpublished). Exon 5 is unique with respect to known MEPE sequences and may itself have some interesting splice variations that are currently being clarified (Fisher et al., unpublished). We have found cDNAs corresponding to mRNA containing exons 4 and 5 together as well as exon 5 alone (Figure 3). We have not yet seen the cDNA corresponding to exon 4 alone like that seen for the monkey, but additional work is being done.

Figure 3 shows the intron-exon structure of the six SIBLING family members. In each case, the full length human mRNA-derived cDNA sequence was compared with the Human Genome Project (www.ncbi.nlm.nih.gov/genome/seq/HsBlast.html) and the exons deciphered from the matches using standard intron donor/acceptor sites. Exon 1 is always a noncoding exon. ENAM is the only member that has a second noncoding exon. The next exon, exon 2 for most members, always contains the start codon, the leader sequence, and the codons for the first two amino acids of the mature proteins. The leader sequence encodes the series of

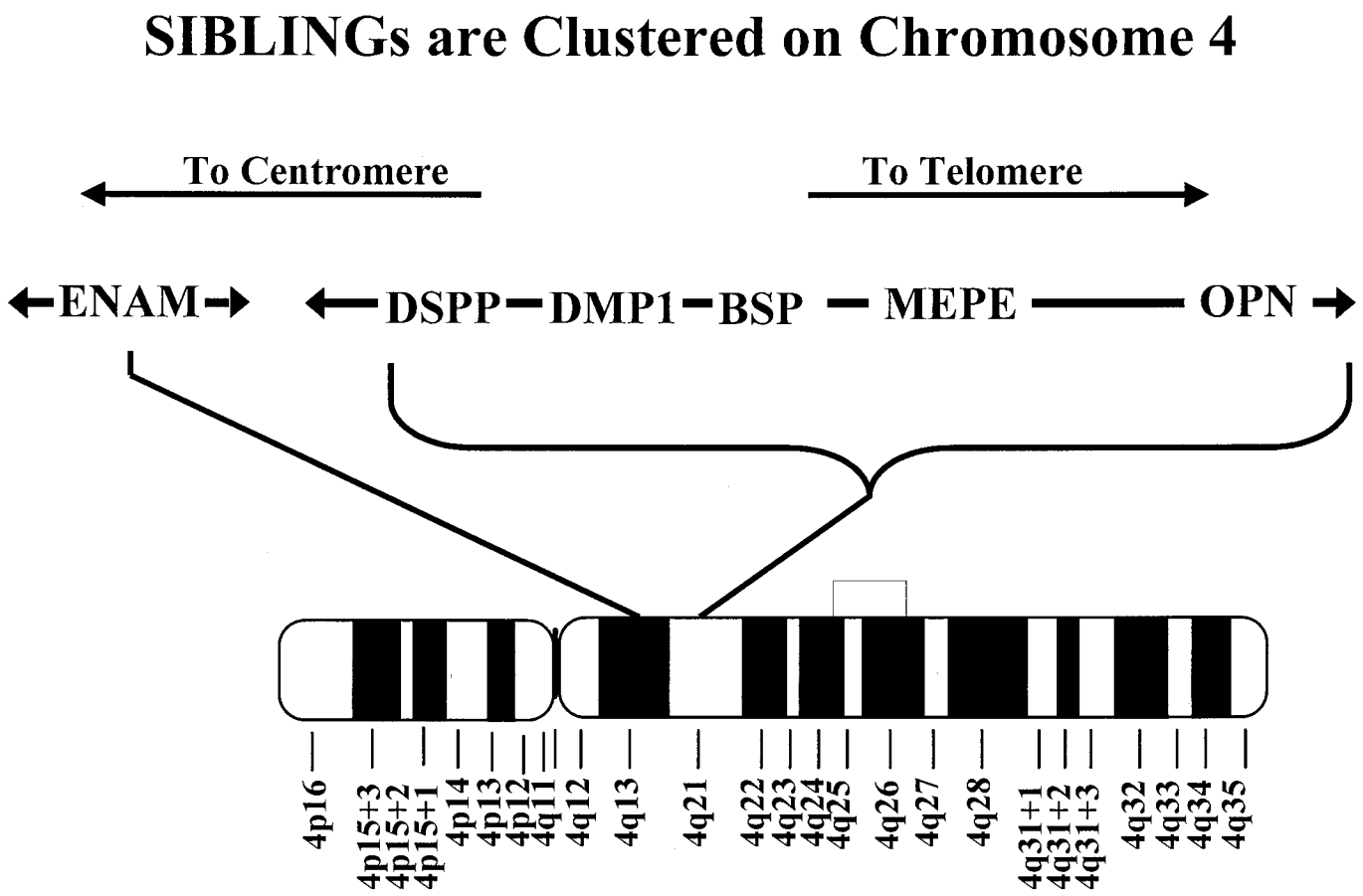

Figure 2. Five SIBLING genes cluster closely together on human chromosome 4 and a sixth candidate gene, ENAM, is located nearer the centromere. DSPP, DMP1, BSP, MEPE, and OPN are located at 4q21.3, within a region of about 750,000 basepairs. ENAM is currently assigned a position of 4q13. These locations are based on the current build (number 26) of the Human Genome Project. The sequencing of this portion of chromosome 4 is incomplete at this time, however, and the final positions may change to some small degree.

hydrophobic amino acids that directs the protein synthesis into the rough endoplasmic reticulum for post-translational modification and subsequent secretion out of the cell. Enamelin's exon 3 does not contain a classical leader sequence and may not be processed and secreted in the same way. Intron 2 and all other introns in the family interrupt the coding sequences between codons (type 0). This implies that any exon can be spliced in or out of the mRNA and not cause a frame shift. Exon 3 usually contains a casein kinase II phosphorylation site (SSEE) and exon 4 is usually relatively proline-rich (PPPP). Exon 5 usually contains another casein kinase II phosphorylation site and, like all the first four exons, is a small exon. The last one or two exons encode the vast majority of the protein (Figure 3 is not drawn to scale) and always contain the integrin-binding tripeptide, RGD. Again ENAM is more distantly related as only the human sequence encodes for the RGD. Pig [13] and mouse [14] do not contain the RGD within their reported sequences.

The exons used in the splice variants are generally conserved. As shown as gray boxes in Figure 3, splice variants missing exon 4 have been reported in OPN [15] and MEPE [5 and this Article). Splice variants missing exon 5 has been shown for OPN [15, 16], DMP1 [3] and MEPE [5 and this article]. To date, there has been no direct proof that the splice variants differ in function.

Two of the SIBLINGs, BSP and OPN, have had their structures solved by NMR. Both were found to be entirely flexible in solution [9]. Flexibility is a common property found in proteins or domains of proteins that have a number of different binding partners. Once bound to their multiple partners, the proteins often have a single conformation. For example, in isolation the protein L39e from the large subunit of ribosomes is completely flexible in solution, but its structure is well defined within the assembled ribosome and can be observed in X-ray diffraction [17]. Other ribosomal proteins (L2, L3, L4, etc.) have large domains that are completely unstructured in solution but also have single fixed structures in the assembled ribosome. None of the closely clustered five SIBLINGs has more than one cysteine within their mature sequences, so there is no chance of forming intramolecular disulfide bonds, but again ENAM is a likely exception. In contrast, most of the structured secreted proteins such as osteonectin, osteocalcin, decorin, biglycan etc. have disulfide bonds to help stabilize their three dimensional shapes in the extracellular environment. Furthermore, while the general

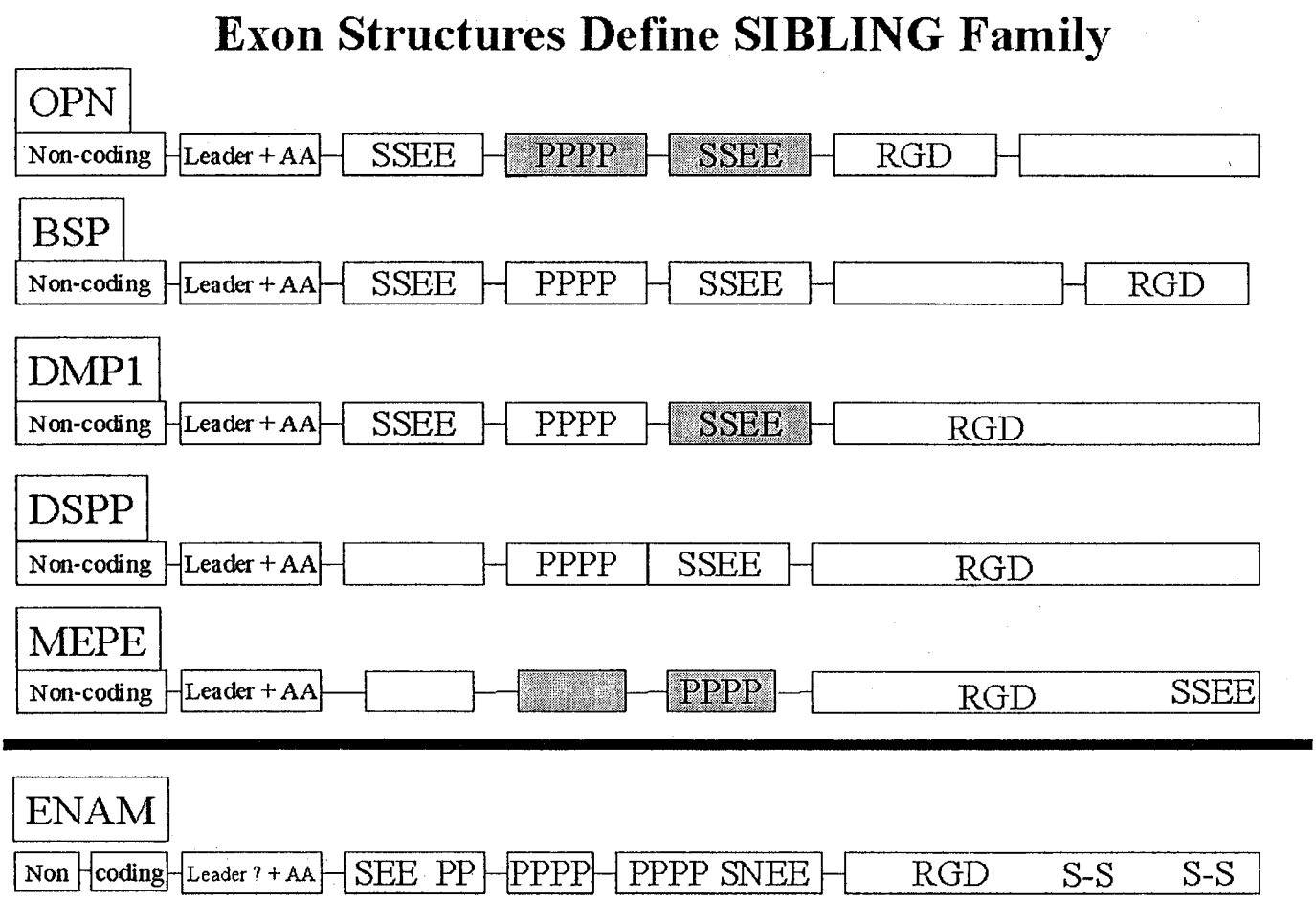

Figure 3. Exon structure defines the SIBLING family. The exon structures of the six candidate genes for the SIBLING family are illustrated. Exons are drawn as boxes and introns as connecting lines. Exon 1 is noncoding. For all but ENAM, exon 2 encodes for the leader sequence plus the first two amino acids of the mature protein. Exon 3 often contains the consensus sequences for casein kinase II phosphorylation (SSEE), as does exon 5. Exon 4 is usually relatively proline rich (PPPP). The last one or two exons encode the vast majority of the protein (figure not drawn to scale) and always contain the integrin-binding tripeptide ArgGlyAsp (RGD). The shadowing of exons illustrates those exons known to be involved in splice variants. ENAM is a more distantly related gene that has two noncoding 5′ and is also likely to contain disulfide bonds (S-S) that the other SIBLINGs do not.

chemical properties of the amino acids along each SIBLING's length are conserved across the animal species (hydrophilic and either acidic or basic, etc.), there is a great deal of divergence within each protein.

All the amino acid sequences of the SIBLINGs are only 55–73% identical between mouse and human while other noncollagenous proteins that are thought to have stable structures in solution are more highly conserved (for example, osteonectin, 96% [18]; biglycan, 91% [19, 20]; and matrix gla protein, 84% [21, 22]. It seems reasonable that proteins that encode for small conserved contact points for a number of binding partners spaced throughout their lengths could have many other regions that can mutate to other amino acids as long as they maintain their hydrophilicity and flexibility in solution. A corollary to this hypothesis is that the short stretches of amino acids that are conserved across species are likely to be directly or indirectly involved in binding other proteins. The tripeptide RGD is one example of the conservation of a short series of amino acids that is involved in binding to other protein complexes, the subfamily of integrins.

Other known binding partners of most of the acidic SIBLINGs include complement Factor H (BSP, DMP1, and OPN) and CD44 (DMP1 and OPN). A short summary of the alternate pathway of complement may be helpful at this point (Figure 4) [23]. The alternate complement pathway (ACP) is one of the most ancient of immune responses, predating the better known pathways that involve specific antibodies. The ACP involves about 20 different proteins and together they constitute approximately 5% of the serum proteins by weight. The proteins in the blood, although found in high concentration, are in conformations that do not favor interactions until an activation cascade is triggered. Briefly, the triggering event is when one component, C3 (found at ~1 mg/ml in the serum), undergoes a spontaneous rearrangement and exposes a highly reactive chemical group. If, within a few milliseconds, this activated C3 can come into contact with a free OH or NH group on a carbohydrate or

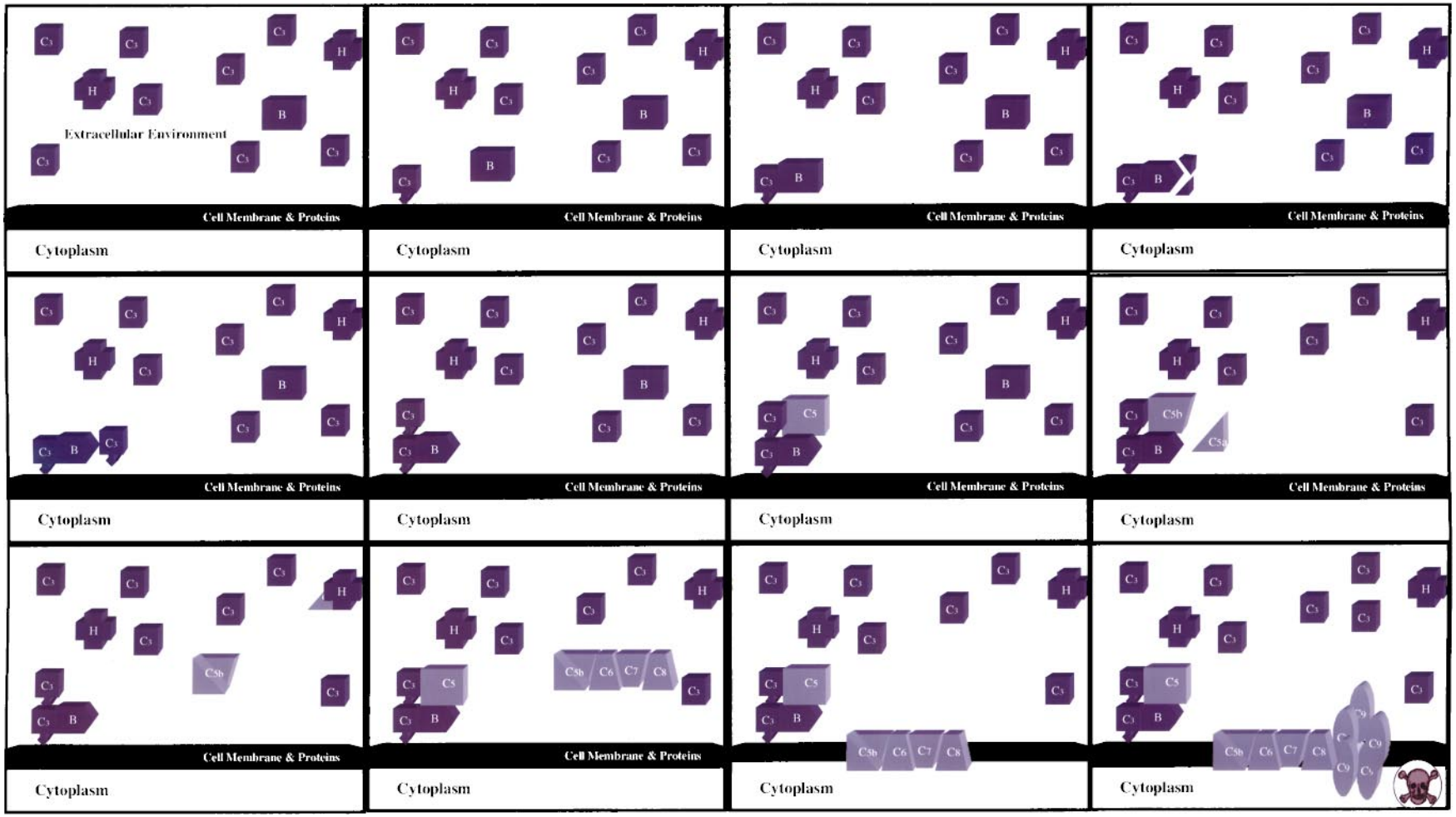

Figure 4. A cartoon description of the basic steps in the lytic process for the alternate complement pathway. See text for a brief description of the process.

protein, then the C3 will form a covalent bond with that molecule. The design is that this reactive carbohydrate or protein be on the surface of an invading bacterium or parasite, but in fact the reaction can and does occur on the surface of all cells. Furthermore, although the C3 spontaneous activation is rare, the number of C3 proteins is sufficiently high that a typical invading bacterium is likely to be bound by one or more C3 molecules within a few minutes. The bound C3 protein is in a conformation that permits the binding of the next complement protein, Factor B. Factor B itself undergoes a conformational change upon binding, and becomes a substrate for serum protease (Factor D). The digested Factor B, now designated Bb, is an activated protease that can bind C3 and cleave it into two pieces, C3b and C3a. (The C3a diffuses away and is a potent chemoattractant for various immune cells.)

The formation of the C3b exposes the chemically reactive group and, due to its proximity to the cell surface, often binds to the same cell. Because the bound C3b (and its various breakdown products) is a ligand for certain receptors on immune cells, this opsonization process can itself lead to the destruction of the labeled cell. However, a second process also occurs. Occasionally the activated C3b molecule forms the covalent bond not with the cells surface but on the C3Bb complex itself. When this happens, another complement protein, C5, can be bound into the complex and is cleaved by the Bb protease into two pieces, C5a and C5b. (The smaller C5a diffuses away and is another very potent immune activating molecule.) The C5b protein is released and now can bind C6. The induced conformation of C6 then binds C7, which then binds C8. Together the C5bC6C7C8 complex inserts into a cellular membrane and then binds a series of C9 proteins from the serum. As the number of C9 proteins increase, a pore forms in the membrane, killing the cell by depolarizing the cell and permitting an exchange of diffusible elements between the inside and outside of the cell. This is called the lytic pathway of the alternate complement process.

Most normal, healthy mammalian cells survive the continuous attack by this pathway by three mechanisms. Many cells can produce two different membrane-associated proteins that can disrupt this lytic pathway of the ACP. Decay accelerating factor (DAF, CD55) has a higher affinity for C3b than does Bb so it can displace the protease and stop the cascade. Another protein, membrane co-factor protein (MCP, CD46), can bind to complement Factor I. This binding causes a conformational change in Factor I and thereby enhances its proteolytic function. The MCP/Factor I complex can then digest the C3b and destroy its ability to promote the lytic cascade. A third mechanism involves a complement protein found at ~0.5 mg/ml in the serum, Factor H. By itself in the blood, Factor H has a real but weak affinity for C3b and can display weak DAF-like activity. Factor H also has a low affinity for Factor I and thereby can act as a poor but measurable cofactor for Factor I and display MCP-like activity. But in both cases, Factor H can itself bind to proteins or carbohydrate groups, undergo a conformational change, and acquire a higher affinity for C3b and/or Factor I. We have shown previously that at least three of the SIBLINGs have the ability to bind Factor H and confer protective activity on cells (*BBRC*, IADR abstract and submitted).

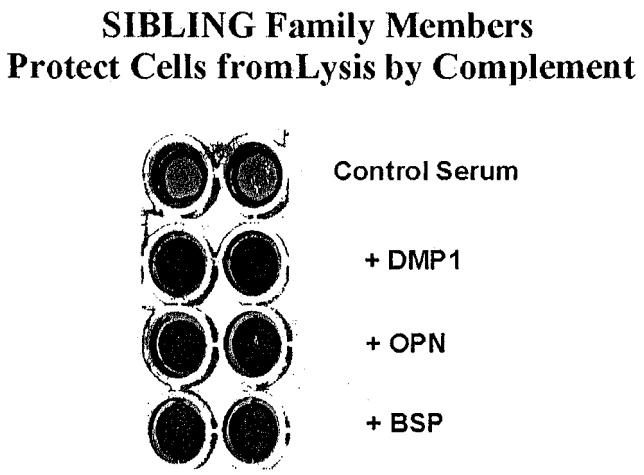

Figure 5. SIBLING family members protect cells from lysis by the alternate complement pathway. Murine erythroleukemia cells in microtiter plate wells are lysed by complement in human serum (control serum). Lysed cells cannot process the clear MTT reagent into the dark blue color and the wells remain clear. Pretreating the cells with DMP1, OPN, or BSP prior to the addition of the human serum protects these cells from the lytic pathway. The living cells process the MTT to the dark color that is seen in each of the SIBLING-treated wells.

When mouse erythroleukemia (MEL) cells are treated with dilute human serum, the cells are lysed by the ACP resulting in cells that cannot metabolize the colorless thiazolyl blue (MTT) to the characteristic blue color (Figure 5). Similar results can be

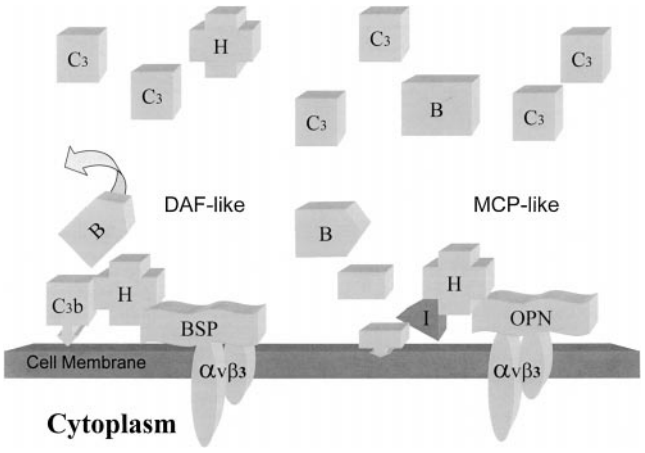

Figure 6. Two possible mechanisms for SIBLINGs to protect cells from alternate complement pathway. BSP, OPN, or DMP1 binds first to an integrin through its RGD domain (or CD44 for OPN and DMP1 but not BSP) and then bind complement Factor H. The Factor H then undergoes a conformational change and either has a higher affinity for C3b than does Bb and displaces the protease (DAF-like activity) or has a higher affinity for Factor I that then can degrade the C3b (MCP-like activity). Interestingly, if the SIBLING binds to Factor H before it encounters a cell surface receptor, the SIBLING cannot bind to the receptor and the activity is not acquired. This limits this biological activity to short distances from the site of the secretion of the SIBLING.

seen by using human cells with guinea pig serum as the source of active complement. When we treat the cells with BSP, DMP1, or OPN, and then expose them to the complement in human serum, the cells are not lysed. These three acid SIBLINGs are stopping the lytic pathway of the ACP. Blocking the ability of the SIBLINGs to bind to their cells surface receptors by mutating the RGD to KAE, using antibodies that block the receptors etc., negates the protective ability of the three SIBLINGs [9, 24, and 25]. This shows that the protective properties of BSP, OPN, and DMP1 all work in conjunction with a cell surface receptor. Furthermore, because we have clearly shown that these proteins form a strong 1:1 complex with complement Factor H, it is reasonable to speculate that the protection provided by the three SIBLINGs is due to a simultaneous complex of Factor H, SIBLING, and cell surface receptors. We have hypothesized that these complexes likely mimic either DAF (displacing the protease, Bb, from the complex) or MCP (acting as a cofactor for Factor I) as drawn in Figure 6. Future studies will determine which of these pathways is the correct model.

CONCLUSION

There are clearly five members of the SIBLING family that cluster together on human chromosome 4, chromosome 5 in the mouse. All these are charged, possibly flexible proteins that may contain little or no secondary structure when isolated in solution but are likely to have structure induced when they interact with one or more of their binding partners. Binding partners for various members of the SIBLING family include cell surface proteins for all of them (integrins, CD44, etc.), Factor H for at least three of them, hydroxyapatite for the acid members, and likely other proteins in the future. ENAM is a more distantly related protein whose gene is found more centrally on the same chromosome and that is likely to contain secondary and tertiary structure.

REFERENCES

[1] Myllyharju, J., and Kivirikko, K.I. (2001). Collagens and collagen-related diseases. *Ann. Med.* 33:7–21.

[2] Fisher, L.W., McBride, O.W., Termine, J.D., and Young, M.F. (1990). Human bone sialoprotein: Deduced protein sequence and chromosomal localization. *J. Biol. Chem.* 265:2347–2351.

[3] Hirst, K.L., Simmons, D., Feng, J., Aplin, H., Dixon, M.J., and MacDougall, M. (1997). Elucidation of the sequence and the genomic organization of the human dentin matrix acidic phosphoprotein 1 (DMP1) gene: Exclusion of the locus from a causative role in the pathogenesis of dentinogenesis imperfecta type II. *Genomics* 42:38–45.

[4] Gu, K., Chang, S.R., Slaven, M.S., Clarkson, B.H., Rutherford, R.B. and Ritchie, H.H. (1997). Human dentin phosphophoryn nucleotide and amino acid sequence. *Eur. J. Oral Sci.* 106:1043–1047.

[5] Rowe, P.S., de Zoysa, P.A., Dong, R., Wang, H.R., White, K.E., Econs, M.J., and Oudet, C.L. (2000). MEPE, a new gene expressed in bone marrow and tumors causing osteomalacia. *Genomics* 67, 54–68.

[6] Petersen, D.N., Tkalcevic, G.T., Mansolf, A.L., Rivera-Gonzalez, R., and Brown, T.A. (2000). Identification of osteoblast/osteocyte factor 45 (OF45), a bone-specific cDNA encoding an RGD-containing protein that is highly expressed in osteoblasts and osteocytes. *J. Biol. Chem.* 17:36172–36180.

[7] Kiefer, M.C., Bauer, D.M., and Barr, P.J. (1989). The cDNA and derived amino acid sequence for human osteopontin. *Nucleic Acids Res.* 17:3306.

[8] Hu, C.C., Hart, T.C., Dupont, B.R., Chen, J.J., Sun, X., Qian, Q., Zhang, C.H., Jiang, H., Mattern, V.L., Wright, J.T., and Simmer, J.P. (2000). Cloning human enamelin cDNA, chromosomal localization, and analysis of expression during tooth development. *J. Dent. Res.* 79:912–919.

[9] Fisher, L.W., Torchia, D.A., Fohr, B., Young, M.F., and Fedarko, N.S. (2001). Flexible structures of SIBLING proteins, bone sialoprotein, and osteopontin. *Biochem. Biophys. Res. Comm.* 280:460–465.

[10] Allikmets, R., Schriml, L.M., Hutchinson, A., Romano-Spica, V., and Dean, M. (1998). A human placenta-specific ATP-binding cassette gene (ABCP) on chromosome 4q22 that is involved in multidrug resistance. *Cancer Res.* 58:5337–5339.

[11] Bianco, P., Fisher, L.W., Young, M.F., Termine, J.D., and Robey, P.G. (1991). Expression of bone sialoprotein (BSP) in developing human tissues. *Calcif. Tiss. Int.* 49:421–426.

[12] Bellahcene, A., and Castronovo, V. (1997). Expression of bone matrix proteins in human breast cancer: Potential roles in microcalcification formation and in the genesis of bone metastases. *Bull. Cancer* 84:17–24.

[13] Fukae, M., Tanabe, T., Murakami, C., Dohi, N., Uchida, T., and Shimizu, M. (1996). Primary structure of the porcine 89-kDa enamelin. *Adv. Dent. Res.* 10:111–118.

[14] Hu, C.-C., Simmer, J.P., Bartlett, J.D., Nanci, A., Qian, Q., Zhang, C., Ryu, O.H., Xue, J., Fukae, M., Uchida, T., and McDougall, M. (1998). Murine enamelin: cDNA and derived protein sequences *Connect. Tiss. Res.* 39:47–61.

[15] Saitoh, Y., Kuratsu, J., Takeshima, H., Yamamoto, S., and Ushio, Y. (1995). Expression of osteopontin in human glioma. Its correlation with the malignancy. *Lab. Invest.* 72:55–63.

[16] Young, M.F., Kerr, J.M., Termine, J.D., Wewer, U.M., Wang, M.G., McBride, O.W., and Fisher, L.W. (1990). cDNA cloning, mRNA distribution and heterogeneity, chromosomal location, and RFLP analysis of human osteopontin (OPN). *Genomics.* 7:491–502.

[17] Ban, N., Nissen, P., Hansen, J., Moore, P.B., and Steitz, T.A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science* 289:905–920.

[18] Lankat-Buttgereit, B., Mann, K., Deutzmann, R., Timpl, R., and Krieg, T. (1988). Cloning and complete amino acid sequences of human and murine basement membrane protein BM-40 (SPARC, osteonectin). *FEBS Lett.* 236:352–356.

[19] Rau, W., Just, W., Vetter, U., and Vogel, W. (1994). A dinucleotide repeat in the mouse biglycan gene (EST) on the X chromosome. *Mamm. Genome* 5:395–396.

[20] Fisher, L.W., Termine, J.D., and Young, M.F. (1989). Deduced protein sequence of bone small proteoglycan I (biglycan) shows homology with proteoglycan II (decorin) and several nonconnective tissue proteins in a variety of species. *J. Biol. Chem.* 264:4571–4576.

[21] Ikeda, T., Yamaguchi, A., Icho, T., Tsuchida, N., and Yoshiki, S. (1991). cDNA and deduced amino acid sequence of mouse matrix gla protein: One of five glutamic acid residues potentially modified to gla is not conserved in the mouse sequence. *J. Bone Miner. Res.* 6:1013–1017.

[22] Cancela, L., Hsieh, C.L., Francke, U., and Price, P.A. (1990). Molecular structure, chromosome assignment, and promoter organization of the human matrix Gla protein gene. *J. Biol. Chem.* 265:15040–15048.

[23] Rother, K., Till, G.O., and Hansch, G.M. (1998). The Complement System New York: Springer.

[24] Fedarko, N.S., Fohr, B., Robey, P.G., Young, M.F., and Fisher, L.W. (2000). Factor H binding to bone sialoprotein and osteopontin enables tumor cell evasion of complement-mediated attack. *J. Biol. Chem.* 275:16666–16672.

[25] Jain, A., Karadag, A., Fohr, B., Fisher, L.W., and Fedarko, N.S. (2002). Three SIBLINGS (Small integrin-binding ligand, n-linked glycoproteins) enhance factor H's cofactor activity enabling MCP-like cellular evasion of complement-mediated attack. *J. Biol. Chem.* 277:13700–13708.